

Geometry-driven quantization for omnidirectional image coding

Francesca De Simone and Pascal Frossard
Signal Processing Laboratory (LTS4)

Ecole Polytechnique Fédérale de Lausanne (EPFL)
Lausanne, Switzerland

Paul Wilkins, Neil Birkbeck, Anil Kokaram

Google Inc
Mountain View, California, USA

Abstract—In this paper we propose a method to adapt the quantization tables of typical block-based transform codecs when the input to the encoder is a panoramic image resulting from equirectangular projection of a spherical image. When the visual content is projected from the panorama to the viewport, a frequency shift is occurring. The quantization can be adapted accordingly: the quantization step sizes that would be optimal to quantize the transform coefficients of the viewport image block, can be used to quantize the coefficients of the panoramic block. As a proof of concept, the proposed quantization strategy has been used in JPEG compression. Results show that a rate reduction up to 2.99% can be achieved for the same perceptual quality of the spherical signal with respect to a standard quantization.

I. INTRODUCTION

From virtual reality to robotics, applications exploiting omnidirectional images and videos are expected to become wide spread in the near future. Cameras able to instantaneously capture the 360° surrounding natural scene have already started to appear as consumer products [1] and professional tools [2]. Omnidirectional imaging involves a particular acquisition and rendering geometry: the image can be considered as a signal lying on a *viewing sphere*, which the viewer, at the center of the sphere, can navigate by changing viewing direction. When the content is rendered to be visualized, a portion of the sphere surface is projected to a planar segment, the *viewport*, depending on the user's viewing direction.

Omnidirectional signals, being intrinsically spherical, differ from those captured with directional cameras and imply higher data rates: strategies to compress omnidirectional images and videos are thus needed and their design is an open research topic. While the design of new coding solutions tuned to spherical signals [3][4] is definitely of interest, encoding omnidirectional content using existing image and video codecs [5][6][7][8][9][10][11] may present clear advantages, since these codecs are integrated in many multimedia pipelines and have been optimized over decades of research efforts.

In order to use standard processing techniques and file formats, the output of fully omnidirectional cameras is usually a rectangular *panoramic image* (i.e., *panorama*), resulting from the projection of the viewing sphere to a plane (*map projection*) [12]. The *equirectangular projection* is the simplest and most used projection, where the longitude and latitude of

each point on the sphere surface are used as coordinates on the plane.

A panoramic image can be given as input to standard image and video encoders but it is not a classical natural image: depending on the map projection, some of the assumptions upon the frequency characteristics of the signal exploited by the encoder, might be not verified.

Particularly, the quantization steps used in most of the existing block-based transform coders to quantize the transform coefficients of the image blocks or their prediction error, have been defined in order to minimize the distortion between the original image and the decoded image taking into account the frequency characteristics of natural images and the human visual sensitivity to different frequency components. In the omnidirectional encoding scenario, the transform coefficients undergoing quantization are those of the panorama, whose statistics are not necessarily those of the transform coefficients of natural images. Also, the image presented to the viewer, on which the distortion should be minimized, is the viewport rather than the panorama. An adaptation of the signal given as input to the encoder or of the quantization strategy is thus needed and expected to be beneficial in terms of rate-distortion encoding performance.

In this paper, we propose a method to adapt the quantization tables used by block-based transform coders, when the input to the encoder is a panoramic image resulting from equirectangular projection. The adaptation is geometry-driven since it varies depending on the position of the block of the panorama when projected on the viewing sphere. The proposed method relies on the following observations. The viewport is a natural image, so the quantization steps defined in the standard codecs would be effective if used to compress the viewport directly. Since each transform coefficient specifies the contribution of a sinusoid pattern at a particular frequency to the actual signal and the projection from the viewport to the panorama is changing the frequencies of the signal, we can model this shift of frequencies in order to define a rule to select the appropriate quantization step for each transform coefficient of the panorama. Equirectangular projection is considered in this paper, but the analysis extends to other projections. As a proof of concept, the proposed quantization strategy has been used in JPEG compression. Results show that a rate reduction

up to 2.99% can be achieved for the same perceptual quality of the spherical signal with respect to standard quantization.

In order to mitigate the influence of geometry distortion, map projections alternative to the equirectangular one have been proposed in [8] and [9]. The impact of different map projections on the performance of a standard coding scheme has been analysed in [13]. These projections do not ensure that the distortion measure on the panoramic image aligns with a distortion measure derived on the viewport and would still benefit from a formulation that relates frequency characteristics on the projection to that in a viewport image, as we propose. Furthermore, the equirectangular projection has advantages as a mezzanine format, in that it has a simple analytic form and can be interpreted without rendering to a viewport.

When the panorama is the result of the equirectangular projection, tile-based panoramic encoding schemes have been proposed [10][11]. In these schemes, the panorama is divided into independently decodable tiles that are assigned with different encoding rates [10] and undergo a subsampling pre-processing step [11]. Nevertheless, the quantization tables defined for standard images are used to quantize the transform coefficients of the tiled panorama, regardless of their optimality. In this work, we rather propose a method to adapt the quantization process.

In Section II, the geometry and projections defining the scenario under analysis are detailed. Section III describes the geometry-driven quantization and its application. Section IV shows the results obtained when using the proposed method in JPEG compression. Section V concludes the paper.

II. FRAMEWORK

A. Viewing sphere and viewing direction: notation

The omnidirectional visual content is generally represented as a spherical image. In particular, an omnidirectional camera, located at the origin of the right-handed world coordinate system, projects the point $\mathbf{X} = (X, Y, Z)^T$ in the 3D space to the point $\mathbf{x} = (x, y, z)^T$ on the spherical imaging surface of radius r , usually considered unitary, i.e., the viewing sphere. In spherical coordinates, each point on the sphere surface is identified by its longitude ($0 \leq \theta < 2\pi$), and its latitude ($0 \leq \phi \leq \pi$). The user is assumed to be at the centre of the sphere and can change his viewing direction to navigate the omnidirectional content. The viewing direction is defined by the elevation ($-\pi/2 \leq el \leq \pi/2$) and the azimuth ($-\pi \leq az < \pi$), with:

$$\begin{aligned} el &= \frac{\pi}{2} - \phi \\ az &= \begin{cases} -\theta & \text{if } 0 \leq \theta \leq \pi \\ \theta - \pi & \text{if } \pi \leq \theta < 2\pi. \end{cases} \end{aligned} \quad (1)$$

B. From the sphere to the panorama

The spherical image is usually projected to a plane, in order to be easily processed. When *equirectangular projection* is applied, each point on the sphere surface is projected onto a plane by using its longitude θ and latitude ϕ as coordinates on

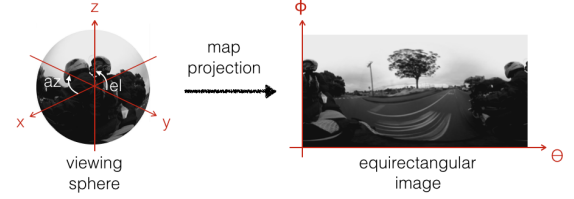


Fig. 1. From the viewing sphere to the equirectangular panorama.

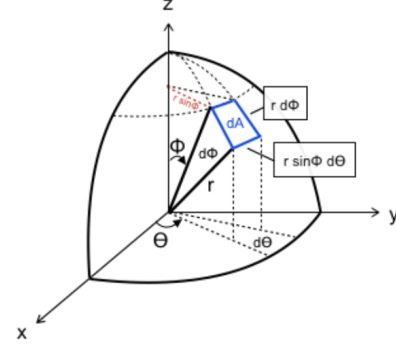


Fig. 2. Area element defined by the grid of meridians and parallels on the sphere: $dA_s = r^2 \sin\phi \, d\phi \, d\theta$.

the plane. Each row and column of the panoramic image can thus be associated to a ϕ and θ value, respectively. The width of the panorama is twice its height ($im_h = im_w/2$). Since the pixel distribution on the panorama is uniform and pixels are square, the projection implies that one pixel is defined on the sphere surface for each (θ, ϕ) and step $(d\theta, d\phi)$, with $d\theta = d\phi = \frac{2\pi}{im_w} = \frac{\pi}{im_h}$.

The visual content, originally captured on a spherical surface, may appear severely distorted on the equirectangular panorama, especially towards the upper and lower parts of the panorama (Fig. 1). This distortion can be easily understood considering a grid of meridians and parallels connecting equally spaced points at step $(d\theta, d\phi)$ on the sphere surface. At high resolution, each area element defined by this grid on the sphere has magnitude $dA_s = r^2 \sin\phi \, d\phi \, d\theta$ (Fig. 2). After equirectangular projection, the area element becomes equal to $dA_p = r^2 \, d\theta \, d\phi$. This implies that the pixel area is stretched by a factor $1/\sin(\phi)$ (i.e., $1/\cos(el)$) from the sphere to the panorama, i.e., pixels towards the poles are stretched more than pixels towards the equator. The distortion is essentially due to the horizontal stretching, since the vertical distance between two consecutive parallels is the same on the sphere and on the panorama.

C. From the sphere to the viewport

A viewport (Fig. 3) is defined by: the viewing direction (el_o, az_o) , which identifies the center O' where the viewport is tangent to the sphere; its resolution $[vp_h, vp_w]$; its horizontal and vertical field of view, FoV_H and FoV_V respectively. Only one of the two fields of view needs to be specified if the aspect ratio of the viewport is fixed.

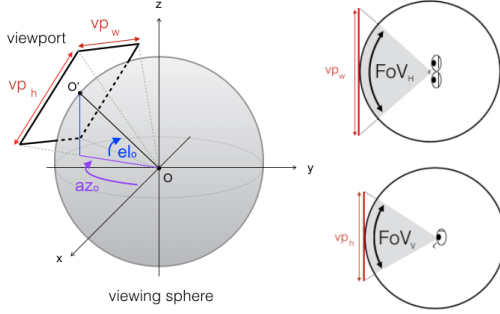


Fig. 3. Parameters defining a viewport.

The projection of points from the sphere surface to a plane tangent to it at any point on the sphere surface is an azimuthal projection, known as *oblique gnomonic projection* [12]. This projection causes a distortion of the visual content from the sphere surface to the viewport that varies depending on the position on the viewport: the more it moves away from the viewport center, the more the visual content is stretched both in horizontal and vertical directions. Also, visual content on a portion of the sphere surface may be projected to different viewports, thus, undergo different distortions.

III. GEOMETRY-DRIVEN QUANTIZATION

In block-based transform coding, each image block is expressed as a linear combination of basic block patterns at different horizontal and vertical frequencies, i.e., the *basis images*. Each transform coefficient represents the contribution from one basis image to the signal. Quantization tables, specifying the step size to be used to quantize each transform coefficient, have been defined in order to minimize the distortion on the compressed signal, taking into account the frequency characteristics of natural images and possibly the human visual sensitivity to different frequency components.

In the scenario considered in this paper, the input to the encoder is the equirectangular panoramic image, while the distortion should be minimized on the rendered planar images, namely the viewports. Due to the projections described at Section II, the signal on the panorama is a warped version of the signal on the viewport: depending on the elevation at which the viewport is extracted, a given horizontal or vertical spatial frequency on the panorama (f_{Hp} and f_{Vp} , respectively) may result in a different frequency on the viewport (f_{Hvp} and f_{Vvp}). Each transform coefficient of the panoramic block, specifying the contribution of a frequency f_{Hp} and f_{Vp} to the signal on the panorama, can thus be quantized using the quantization step defined for the corresponding frequency f_{Hvp} and f_{Vvp} on the viewport.

The model of frequency shift adopted and the corresponding shift in terms of Discrete Cosine Transform (DCT) basis images are described in Section III-A. Details of the implementation of the geometry-driven quantization are provided in Section III-B.

A. Approximation for frequency shift

In order to illustrate the impact of the projections described at Section II on the frequency characteristics of the visual content that undergoes compression, we consider an example with simple artificial patterns. Since the input of our encoding pipeline is the equirectangular panoramic image, we start our analysis from it.

A 1D sinusoidal wave at horizontal frequency f_{Hp} (i.e. a vertical pattern) on a pixel row of the panorama corresponding to elevation el , becomes a sinusoidal signal at horizontal frequency f_{Hs} when projected on the sphere via inverse equirectangular projection. Due to the geometry of the equirectangular projection, $f_{Hp} \leq f_{Hs}$, since:

$$f_{Hs} = f_{Hp} / \cos(el). \quad (2)$$

In other words, the more we move from the central row of the panoramic image towards the upper or lower rows, the more the horizontal spatial frequencies increase when projected from the panorama to the sphere. We will refer to this shift as *horizontal frequency shift*.

On the contrary, when a horizontal pattern is projected from the equirectangular image to the sphere there is no change in terms of frequencies, i.e. there is no vertical frequency shift.

As mentioned in Section II-C, the effect of the oblique gnomonic projection depends on where the projected point is with respect to the center of the viewport. In terms of frequencies, this implies that the sinusoidal signal on the sphere at frequency f_{Hs} and elevation el (or frequency f_{Vs} and azimuth az) becomes a wave whose horizontal frequency f_{Hvp} (or vertical frequency f_{Vvp}) is varying as a function of the horizontal (or vertical) displacement with respect to the viewport center, in the viewport. Additionally, each point on the sphere can be projected to different viewports, thus have a different displacement with respect to the center of different viewports. While the exact mapping of frequency from the panoramic image to a specific viewport can be derived, its use for our scope is impractical when the panoramic image is compressed considering all viewing directions equally likely, i.e. each point on the sphere may project to any point on the viewport of some gnomonic projection. This is similar to integrating over all possible viewing directions, meaning the effect of the gnomonic projection is negligible. Thus, we approximate the ideal frequency shift from the panorama to the viewport, affecting both horizontal and vertical frequencies, as the horizontal frequency shift at (2), occurring when the panoramic image is projected to the sphere.

If the contribution of any frequency to the image block is expressed in terms of DCT basis images, the shift at (2) becomes a shift from the (k' -th, l -th) basis image on the panorama to the (k -th, l -th) basis image on the sphere, with:

$$k = \lceil k' / \cos(el) \rceil \quad (3)$$

where $\lceil \cdot \rceil$ is the ceiling operator, since k and k' can have discrete integer values only.

TABLE I

SHIFT OF HORIZONTAL BASIS VECTOR FREQUENCY INDEX DEPENDING ON ELEVATION ($-\pi/2 \leq el \leq \pi/2$): k' IS THE INDEX ON THE PANORAMA, k IS THE CORRESPONDING INDEX ON THE VIEWPORT.

$k'=$	0	1	2	3	4	5	6	7	el
$k=$	0	7	7	7	7	7	7	7	$[\pm 1.4268, \pm 1.5708]$
$k=$	0	6	7	7	7	7	7	7	$[\pm 1.4006, \pm 1.4268]$
$k=$	0	5	7	7	7	7	7	7	$[\pm 1.3483, \pm 1.4006]$
$k=$	0	4	7	7	7	7	7	7	$[\pm 1.2828, \pm 1.3483]$
$k=$	0	3	7	7	7	7	7	7	$[\pm 1.2697, \pm 1.2828]$
$k=$	0	3	6	7	7	7	7	7	$[\pm 1.2043, \pm 1.2697]$
$k=$	0	3	5	7	7	7	7	7	$[\pm 1.165, \pm 1.2043]$
$k=$	0	2	5	7	7	7	7	7	$[\pm 1.1126, \pm 1.165]$
$k=$	0	2	4	7	7	7	7	7	$[\pm 1.0996, \pm 1.1126]$
$k=$	0	2	4	6	7	7	7	7	$[\pm 0.9948, \pm 1.0996]$
$k=$	0	2	4	5	7	7	7	7	$[\pm 0.9687, \pm 0.9948]$
$k=$	0	2	3	5	7	7	7	7	$[\pm 0.9163, \pm 0.9687]$
$k=$	0	2	3	5	6	7	7	7	$[\pm 0.8508, \pm 0.9163]$
$k=$	0	1	3	4	6	7	7	7	$[\pm 0.7592, \pm 0.8508]$
$k=$	0	1	3	4	5	7	7	7	$[\pm 0.6938, \pm 0.7592]$
$k=$	0	1	3	4	5	6	7	7	$[\pm 0.6545, \pm 0.6938]$
$k=$	0	1	2	4	5	6	7	7	$[\pm 0.5498, \pm 0.6545]$
$k=$	0	1	2	3	5	6	7	7	$[\pm 0.4843, \pm 0.5498]$
$k=$	0	1	2	3	4	6	7	7	$[\pm 0.432, \pm 0.4843]$
$k=$	0	1	2	3	4	5	7	7	$[\pm 0.4058, \pm 0.432]$
$k=$	0	1	2	3	4	5	6	7	$[0, \pm 0.4058]$

B. Adaptation of quantization tables

Each block on the panorama can be associated to an elevation angle, corresponding to the elevation of the center of the block projected on the sphere. Thus, the horizontal frequency shift at (3) can be used to select the quantization step for each DCT coefficient of each block of the panoramic image: the quantization step $\delta_{k,l}$, optimal for the (k -th, l -th) coefficient of the viewport image block, is used to quantize the (k' -th, l -th) coefficient of the block of the panorama at elevation el (*horizontal quantization step size shift*).

Due to the rounding in (3), the shift varies only for ranges of elevation angles: Table I shows the mapping between the frequency index of the sphere and the panorama for each elevation range, considering a panoramic image height of 1920 pixels. Moving towards the poles (i.e., $el = \pm 1.5708$), most horizontal frequencies in the panorama correspond to the highest horizontal frequency on the sphere (i.e. viewport).

IV. EXPERIMENTAL RESULTS

The proposed approach has been tested in a JPEG compression algorithm. The quantization table used in standard JPEG compression is reported in Table II: for example, for a block at $el = \pi/4$, applying the frequency shift rule defined in Table I, the quantization table becomes that reported in Table III. We refer to the JPEG compression of omnidirectional content with the proposed quantization table shift rule, as JPEG-360.

Table IV describes the six equirectangular panoramic frames used to produce the results presented in this section. The frames have been compressed using standard JPEG and JPEG-360, at quality factors in the range [10, 80], with step 5.

The quality of the compressed omnidirectional images has been quantified by means of:

- the focus-of-attention (FoA) weighted spherical-PSNR (W-S-PSNR) [13], computed between the original

TABLE II

JPEG 8×8 QUANTIZATION MATRIX DEFINED IN ANNEX K OF [14].

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

TABLE III

8×8 QUANTIZATION MATRIX USED TO QUANTIZE DCT COEFFICIENTS OF A BLOCK OF THE PANORAMIC IMAGE AT $el = \pi/4$, WHEN THE PROPOSED ADAPTATION IS APPLIED.

16	11	16	24	51	61	61	61
12	12	19	26	60	55	55	55
14	13	24	40	69	56	56	56
14	17	29	51	80	62	62	62
18	22	56	68	103	77	77	77
24	35	64	81	113	92	92	92
49	64	87	103	120	101	101	101
72	92	98	112	103	99	99	99

panorama and the panorama compressed with JPEG and JPEG-360, on a set of uniformly distributed points on the sphere, equal to one fourth of the total number of pixels in the panoramic image;

- the PSNR computed between the viewports extracted from the original panorama and those extracted from the panorama compressed with JPEG and JPEG-360. Nine viewports have been analysed, with $FoV_V = 65^\circ$, $[vp_w, vp_h] = [640, 480]$ pixels, $az_o = 0$ and $el_o = \{\pm\pi/2, \pm 3\pi/8, \pm\pi/4, \pm\pi/8, 0\}$.

The rate is computed on the compressed panoramic images. The Bjontegaard gain (Δ_{BD}) [15] in terms of average percentage of rate difference for the same objective quality is reported in Tables V and VI, for the W-S-PSNR and the viewport PSNR, respectively. The proposed method has better performance than the benchmark when Δ_{BD} is negative, meaning that for the same objective quality a rate saving is reached when compressing the panoramic image.

The results show that an average rate saving is always achieved if the quality of the spherical image is considered. When the quality is computed on viewport images extracted from the sphere, the gain varies depending on the content and the elevation of the viewport. Particularly, for the same content, the gain is not symmetric with respect to the equator

TABLE IV
RESOLUTION OF EQUIRECTANGULAR TEST FRAMES.

Content	Scene description	im_h	im_w
Motocross	Motocross competition.	960	1920
Motorbike	Two people riding a motorbike.	960	1920
Rollercoaster	Two people on a rollercoaster.	1024	2048
Seal	Seals and instructors in a swimming pool.	1024	2048
Motherboard	Synthetic scene depicting a room with electronic equipments.	1448	2898
Cars	Two cars in an autodrome.	1920	3840

TABLE V

Δ_{BD} IN TERMS OF AVERAGE PERCENTAGE OF BITRATE DIFFERENCE BETWEEN JPEG-360 AND JPEG FOR SAME W-S-PSNR OF COMPRESSED PANORAMA: JPEG-360 HAS BETTER PERFORMANCE THAN BENCHMARK WHEN $\Delta_{BD} < 0$.

content	Δ_{BD} rate
Cars	-0.92
Motherboard	-2.99
Motocross	-0.21
Motorbike	-0.92
Rollercoaster	-1.02
Seal	-0.76

TABLE VI

Δ_{BD} IN TERMS OF AVERAGE PERCENTAGE OF BITRATE DIFFERENCE BETWEEN JPEG-360 AND JPEG FOR SAME PSNR OF VIEWPORT EXTRACTED FROM COMPRESSED PANORAMA: JPEG-360 HAS BETTER PERFORMANCE THAN BENCHMARK WHEN $\Delta_{BD} < 0$.

content	$el_o =$	$-\pi/2$	$-3\pi/8$	$-\pi/4$	$-\pi/8$	0	$\pi/8$	$\pi/4$	$3\pi/8$	$\pi/2$
Cars		5.78	3.43	2.04	0.09	-2.78	-2.79	-2.71	-2.68	-2.30
Motherboard		20.00	18.62	-0.42	-6.7	-9.07	-5.80	-4.41	18.69	22.88
Motocross		15.63	11.65	6.67	2.27	-3.34	-3.45	-3.24	-3.00	-3.01
Motorbike		7.00	4.82	3.45	-0.87	-2.07	-1.87	-1.56	-0.53	2.72
Rollercoaster		12.25	8.66	2.28	-0.25	-2.18	-2.45	-2.41	-2.32	-2.38
Seal		11.73	8.95	3.76	-1.28	-2.42	-2.34	-1.86	-2.71	-2.88

and, in some cases, there is no gain when using the proposed method. The asymmetry can be explained by the characteristics of the visual content, which are not symmetric with respect to the equator, as can be noticed in Fig. 4, where the spatial complexity of each viewport, quantified as mean standard deviation of pixel values, is depicted. While the proposed method results in a gain on most of the viewports for $-\pi/8 \leq el_o \leq \pi/2$, the lack of gain observed across all content on the viewports for $-\pi/2 \leq el_o \leq -\pi/4$ could be explained by the fact that when the content is towards the poles and is textured, the horizontal shift approximation and the resulting quantization applied is too rough. It must be mentioned that, by visual inspection, the perceptual quality of the viewport images appears to be transparent with respect to the reference, even if the results show a loss in terms of PSNR when using JPEG-360.

Finally, while the results obtained by using the proposed method within JPEG compression have been discussed, the proposed solution is generic and can be applied to any block-based encoding method using frequency-dependent quantization.

V. CONCLUSIONS

We have presented a method to adapt the quantization tables of typical block-based transform codecs when the input to the encoder is a panoramic image resulting from equirectangular projection, based on a model of the frequency shift from the panorama to the sphere. The proposed quantization has been illustrated within JPEG compression, showing that rate reduction up to 2.99% can be achieved for the same perceptual quality of the spherical signal with respect to standard quantization.

REFERENCES

- [1] [Online]. Available: <http://ricoh-pentax.com.sg/camera/omni-directional.html>

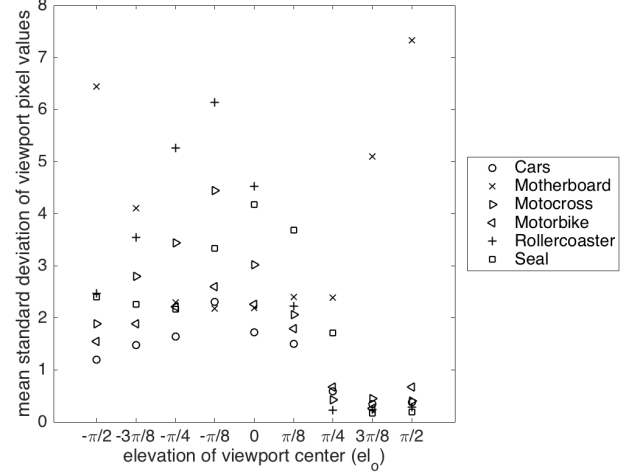


Fig. 4. Spatial complexity of the viewports considered in Table VI.

- [2] [Online]. Available: <https://vr.google.com/jump/>
- [3] I. Tosić, I. Bogdanova, P. Frossard, and P. Vadergheynst, "Multiresolution motion estimation for omnidirectional images," in *European Signal Processing Conference*, 2005.
- [4] I. Tosić and P. Frossard, "Low bit-rate compression of omnidirectional images," in *Picture Coding Symposium*, May 2009, pp. 1–4.
- [5] C. Grunheit, A. Smolic, and T. Wiegand, "Efficient representation and interactive streaming of high-resolution panoramic views," in *International Conference on Image Processing*, vol. 3, 2002, pp. 209–212.
- [6] I. Bauermann, "H.264 based coding of omnidirectional video," in *International Conference on Computer Vision and Graphics*, 2004.
- [7] K.-T. Ng, S.-C. Chan, and H.-Y. Shum, "Data compression and transmission aspects of panoramic videos," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 15, no. 1, pp. 82–95, Jan 2005.
- [8] B. Lelescu, "Representation and coding of panoramic and omnidirectional images," USA Patent US 7,292,722 B2, 2007.
- [9] C.-W. Fu, L. Wan, T.-T. Wong, and C.-S. Leung, "The rhombic dodecahedron map: An efficient scheme for encoding panoramic video," *Multimedia, IEEE Transactions on*, vol. 11, no. 4, pp. 634–644, June 2009.
- [10] P. R. Alfance, J.-F. Macq, and N. Verzijs, "Interactive omnidirectional video delivery: A bandwidth-effective approach," Bell Labs, Tech. Rep., 2012.
- [11] M. Yu, H. Lakshman, and B. Girod, "Content adaptive representations of omnidirectional videos for cinematic virtual reality," in *3rd International Workshop on Immersive Media Experiences. ACM*, 2015, p. 16.
- [12] F. Pearson, *Map Projections Theory and Applications*.
- [13] M. Yu, H. Lakshman, and B. Girod, "A framework to evaluate omnidirectional video coding schemes," in *IEEE International Symposium on Mixed and Augmented Reality*, 2015.
- [14] *Digital compression and coding of continuous still images*, ISO/IEC IS 10918-1 ITU-T Std. Recommendation T.81.
- [15] G. Bjontegaard, "Calculation of average PSNR differences between rd curves," ITU-T VCEG M33, April 2001.